

SETHU S

DATA SCIENTIST | DATA ANALYST

| sethus4791@gmail.com | www.linkedin.com/in/sethus4791 | <https://www.kaggle.com/sethu123123> |
| <https://github.com/itsmesethus> | **PH No:** +91 8122815260 |
| **Address:** No: 2/32, South Street, Anangur, Villupuram, Tamil Nadu - 605 103 |

PROFESSIONAL SUMMARY

Accomplished M.Sc. graduate in Statistics with a proven ability to transform intricate data into actionable insights. Proficient in Statistical Analysis, Hypothesis Testing, Design of Experiments, and Data Modeling, along with expertise in Python, SQL, and R. Known for leading impactful projects and aspiring to drive innovation as a Data Scientist. Committed to continuous growth through active engagement in online courses and collaborations, positioned to elevate your team's success with data-driven solutions

WORK EXPERIENCE

EVOSCIEN UK STATISTICIAN

BANGALORE, INDIA
May 2023 to Present

- Achieved an impressive 30% increase in data-driven decision accuracy by spearheading data acquisition, consolidating datasets, and applying advanced statistical methodologies.
- Influenced Entomological Research direction by designing and executing innovative Hypothesis Tests, leading to the validation of critical project hypotheses.
- Enhanced project outcomes by collaborating with Entomology teams to refine Experimental Design protocols, ensuring high-quality data collection and precise interpretation.
- Transformed complex Entomological data into actionable patterns, producing compelling Data Visualizations. Delivered concise reports, optimizing project success with a 25% efficiency increase.

PROJECTS

MULTICLASS DISEASE CLASSIFICATION OF MEDICAL IMAGE DATA USING CNN (*Jan 2023 – Mar 2023*)

- Developed and fine-tuned convolutional neural networks (CNNs) using TensorFlow and Keras, resulting in an outstanding 94.88% accuracy in classifying Pneumonia X-Ray images.
- Demonstrated exceptional proficiency in multi-class classification by achieving remarkable accuracies: 95.67% for Potato Plant Leaf Disease (3-class), 95.18% for Corn Plant Leaf Disease (4-class), and 90.87% for Gastrointestinal Disease in Humans (8-class).
- Leveraged advanced techniques by merging the VGG19 architecture with strategic layers to boost model performance in the challenging 8-class classification task. Employed Stochastic Gradient Descent to optimize convergence and enhance results.

BCG DATA SCIENCE VIRTUAL INTERNSHIP - "CHURN BUSTER: PREDICTIVE MODELING FOR CUSTOMER RETENTION" (*Aug 2023 – Sep 2023*)

- Drove diverse predictive model implementation, including ANN (Logistic Regression), CatBoost, XGBoost, GradientBoost, and KNeighbors Classifiers, achieving an exceptional 95%+ F1 score with balanced accuracy.
- Strategically engineered data, yielding a 30% prediction error reduction through advanced techniques like SMOTE and Yeo-Johnson transformation, ensuring predictive modeling readiness.
- Converted machine learning insights into effective strategies, resulting in significant customer retention gains and cost savings. Additionally, created impactful Power BI dashboards for data visualization.

QUANTUM DATA ANALYTICS VIRTUAL INTERNSHIP - STORE PERFORMANCE ANALYSIS AND INSIGHTS (*May 2023 – Jun 2023*)

- Reengineered store data analysis for 272 locations, elevating precision with data validation and dynamic visualization. Unearthed a 15% sales surge through meticulous trial and control store assessments.
- Strategically partitioned stores utilizing advanced statistical methods, unmasking a 12% performance differential. Deployed Correlation Analysis and Two-Way ANOVA to pinpoint key drivers.

- Empowered Category Manager with a data-driven report, catalyzing strategic decisions. Guided actions led to a 10% sales boost and optimization of store performance.

PWC SWITZERLAND ANALYTICS VIRTUAL INTERNSHIP - VIRTUAL CASE EXPERIENCES IN TELECOM ANALYTICS (*Jun 2022 – Jul 2022*)

- Designed immersive Power BI dashboards highlighting telecom KPIs such as customer satisfaction, call efficiency, and agent performance (Call Centre Trends).
- Crafted a comprehensive retention KPI dashboard, showcasing data validation, visualization, and wrangling expertise (Customer Retention).
- Led data-driven efforts to promote gender diversity in executive roles, visualizing key diversity KPIs and addressing foundational challenges (Diversity & Inclusion).

EDUCATION

Bharathiar University

Master Of Science in Statistics- CGPA: 8.61

Coimbatore, Tamil Nadu

Sept 2021 – May 2023

- Relevant coursework: Descriptive Statistics, Sampling Theory, Probability Theory, Statistical Estimation Theory, Statistical Quality Control, Multivariate Statistical Analysis, Econometrics, Statistical Inference, Programming in R, Design of Experiments and Stochastic Processes.

Arignar Anna Governments Arts College

Bachelor of Science in Statistics - CGPA: 9.39

Villupuram, Tamil Nadu

Jun 2018 – May 2021

SKILLS & INTERESTS

- **PROGRAMMING LANGUAGES:** Python, R and SQL.
- **LIBRARIES:** Pandas, Numpy, Matplotlib, Seaborn, Plotly, Scipy, Sckit-Learn, TensorFlow, Keras, Statsmodels, Imblearn and Pingouin.
- **TOOLS:** Power BI, IBM SPSS, MINITAB, STATISTICA, Microsoft Excel, MYSQL.
- **DATA PREPROCESSING AND OTHERS:** Data Cleaning, Data Analytics, Exploratory Data Analysis (EDA), Feature Engineering, Feature Selection Techniques, Data Visualization, Outlier Detection, Correlation Analysis, Statistical Tests and etc.
- **ML EXPERTISE (Supervised and Unsupervised ML):** Regression, Classification, Clustering, Predictive Modelling, Quantitative Analysis, Statistical Modelling, CNN.

INTERESTS: Machine Learning, Reading Blogs, Listening to music.

SOFT SKILLS: Goal setting, Research Skills, Decision-making, Active listening, Optimism, Collaboration, Mind-mapping, Networking and Critical observation.

CERTIFICATIONS & WORKSHOPS

- IBM Data Science Professional Certificate – (IBM-Coursera, Jan 2023 – Jul 2023).
- Supervised Machine Learning: Regression And Classification – (Deeplearning.Ai - Coursera, Aug 2023)
- Advanced Learning Algorithms – (Deeplearning.Ai - Coursera, Sep 2023)
- Ensemble Methods In Python – (Datacamp, Sep 2023)
- Feature Engineering For Machine Learning In Python – (Datacamp, Sep 2023)
- Data Analytics With Python – (NPTEL, Swayam, Jan 2022- Apr 2022).
- Databases And SQL For Data Science With Python - (IBM - Coursera, Sep 2022 – Nov 2022).
- Statistical Sciences And Its Application Using Python – (RUSA 2.0 - BCTRC, Bharathiar University, Aug 2022).

PROFILES

- CREDLY: <https://www.credly.com/users/sethu-s.89a25989>
- COURSERA: <https://www.coursera.org/user/b36f2166c75868bc0043c35116bf02a6>